DOCUMENT RESUME

ED 411 245 TM 027 010

AUTHOR Weerasinghe, Dash; Orsak, Timothy; Mendro, Robert
TITLE Value Added Productivity Indicators: A Statistical Comparison of the Pre-Test/Post-Test Model and Gain Model.
PUB DATE 1997-01-00
NOTE 10p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 1997).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; *Achievement Gains; Comparative Analysis; Elementary Secondary Education; Models; *Pretests Posttests; *Productivity; Public Schools; School Effectiveness; Statistical Analysis
IDENTIFIERS Educational Indicators; *Gain Scores; Hierarchical Linear Modeling; *Value Added Model

ABSTRACT

In an age of student accountability, public school systems must find procedures for identifying effective schools, classrooms, and teachers that help students continue to learn academically. As a result, researchers have been modeling schools and classrooms to calculate productivity indicators that will withstand not only statistical review but political criticisms. There are two main approaches to this problem; one uses pretest and posttest scores of students and the other uses student gain scores. In this paper, a model is developed for each approach, and their results are discussed. The outcomes of each model are School Effective Indices (SEI) and Classroom Effective Indices. A set of criteria is established and the SEIs these models produce are tested against these criteria using hierarchical linear modeling (HLM) to determine the "better" approach. It is suggested that both models produce very similar results. If HLM software from SSI Inc. is used, then the pretest-posttest two-level model is more convenient and efficient to use than the three-level Gain model. (Contains four tables and six references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 411 245

Value Added Productivity Indicators: A Statistical Comparison of The Pre-Test/Post-Test Model and Gain Model *

Dash Weerasinghe Timothy Orsak Robert Mendro

Dallas Public Schools
Dallas, TX 75204

March 21, 1997

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Dash Weerasinghe
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract

In an age of student accountability, public school systems must find procedures for identifying effective schools, classrooms and teachers that help students continue to learn academically. As a result researchers have been modeling schools and classrooms to calculate productivity indicators that will withstand not only statistical review but political criticisms. There are two main approaches to this problem; one using Pre-test and Post-test scores of students while the other uses student gain scores. In this paper a model is developed for each approach and their results are discussed. The outcomes of each model are School Effective Indices(SEI) and Classroom Effective Indices(CEI). A set of critera is established and the SEI's the two models are tested against these critera to determine the "better" approach.

Introduction

The strategy for predicting performance from a pre-test score with adjustments for co-variates is becoming the standard for producing productivity indicators. Currently the most asked question is: should we be modeling the student's post-test score or student growth, i.e., gain.

To model student growth or student achievement within a year, the standard statistical approach is to model an outcome variable measuring student's achievement or growth in a

*Paper presented at the annual meeting of the Soutwest Educational Research Association, Austin, TX, January, 1997.

TM027010

BEST COPY AVAILABLE

regression model as follows: If ITBS96_{ij} is the i^{th} student's score on ITBS Reading in school j for year 1995/96, then

$$\text{ITBS96}_{ij} = \beta_0 + r_{ij} \tag{1}$$

where r_{ij} are the student residuals which are $N(0, \sigma)$.

In this model, r_{ij} measures the distance a student's actual score varies from the regression line. If $r_{ij} > 0$, the student performed higher than average and if $r_{ij} < 0$, the student performed lower than average. By aggregating the r_{ij} by each j, we can calculate a measure of how well school j performed. The question of interest is what does r_{ij} measure. We are interested in measuring student achievement in the year 1995/96, but in this particular model r_{ij} is not a measure of student growth, but current standing. This can be easily observed by the bias of r_{ij} with respect to student's previous test scores.

The next progression in this model is to remove any biases in r_{ij} with respect to students' prior test scores, namely ITBS Reading scores from school year 1994/95. Hence equation 1 is expanded as follows:

$$\text{ITBS96}_{ij} = \beta_0 + \beta_1 \text{ITBS95}_{ij} + r_{ij} \tag{2}$$

Now, r_{ij} in the model is conditioned on student's pre-test score and can be said to be unbiased with respect to the students 1994/95 ITBS reading score. Now, r_{ij} is a measure of student growth in year 19945/96.

The next major question arises at this stage: what is the outcome variable that should be modeled? Should it be the student's ITBS Reading score for 1995/96 or should it be the gain in the student's ITBS Reading score from 1994/95 to 1995/96? To answer this question, two types of models will be analyzed, the one used by a large urban school district and a complex gain model recently developed(Bryk and Thum, April 1996). The Bryk and Thum gain model is as follows:

$$\text{GAIN9596}_{ij} = \beta_0 + \beta_1 \text{ITBS95}_{ij} + r_{ij} \tag{3}$$

For this study, the measure GAIN9596 will be calculated as follows:

$$\text{GAIN9596}_{ij} = \text{ITBS96GE}_{ij} - \text{ITBS95GE}_{ij} \tag{4}$$

In this study data from 5197 sixth grade students from 118 elementary schools were modeled, with no missing data.

Methodology

Hierarchical Linear Modeling will be used to compare the Post-Test/Pre-Test Model and the Gain Model. This comparison will be carried out for four models of varying complexities. The school rankings obtained from the two types of models will be compared to each other to observe any differences.

The four models are as follows:

3 BEST COPY AVAILABLE

The basic models with no student level variables and no school level variables:

PP-1 Model:

$$\text{READ96}_{ij} = \beta_{0j} + \beta_{1j}\text{READ95}_{ij} + r_{ij} \tag{5}$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{6}$$
$$\beta_{1j} = \gamma_{10} \tag{7}$$

Gain-1 Model:

$$\text{READGAIN9596}_{ijk} = \pi_{0jk} + \pi_{1jk}\text{READ95}_{ijk} + e_{ijk} \tag{8}$$

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \tag{9}$$
$$\pi_{1jk} = \beta_{10k} + r_{1jk} \tag{10}$$

$$\beta_{00k} = \gamma_{000} + u_{00k} \tag{11}$$
$$\beta_{10k} = \gamma_{100} \tag{12}$$

The basic models with no student level variables and one school level variable, percentage of minority students in a school, PMINORITY:

PP-2 Model:

$$\text{READ96}_{ij} = \beta_{0j} + \beta_{1j}\text{READ95}_{ij} + r_{ij} \tag{13}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{PMINORITY}_j + u_{0j} \tag{14}$$
$$\beta_{1j} = \gamma_{10} \tag{15}$$

Gain-2 Model:

$$\text{READGAIN9596}_{ijk} = \pi_{0jk} + \pi_{1jk}\text{READ95}_{ijk} + e_{ijk} \tag{16}$$

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \tag{17}$$
$$\pi_{1jk} = \beta_{10k} + r_{1jk} \tag{18}$$

$$\beta_{00k} = \gamma_{000} + \gamma_{01k}\text{PMINORITY}_k + u_{00k} \tag{19}$$
$$\beta_{10k} = \gamma_{100} \tag{20}$$

BEST COPY AVAILABLE

Models with student level variables (GENDER, BLACK and HISPANIC) and no school level variables:

PP-3 Model:

$$\begin{aligned}\text{READ96}_{ij} &= \beta_{0j} + \beta_{1j}\text{READ95}_{ij} + \beta_{2j}\text{SEX}_{ij} + \beta_{3j}\text{BLACK}_{ij} + \beta_{4j}\text{HISPANIC}_{ij} \\ &\quad + \beta_{5j}\text{SEX}_{ij} \cdot \text{READ95}_{ij} + \beta_{6j}\text{BLACK}_{ij} \cdot \text{READ95}_{ij} \\ &\quad + \beta_{7j}\text{HISPANIC}_{ij} \cdot \text{READ95}_{ij} + r_{ij} \qquad (21)\end{aligned}$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (22)$$

$$\beta_{qj} = \gamma_{q0} \qquad (23)$$

for $q = 1, 2, \ldots, 7$.

Gain-3 Model:

$$\text{READGAIN9596}_{ijk} = \pi_{0jk} + \pi_{1jk}\text{READ95}_{ijk} + e_{ijk} \qquad (24)$$

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}\text{SEX}_{jk} + \beta_{02k}\text{BLACK}_{jk} + \beta_{03k}\text{HISPANIC}_{jk} + r_{0jk} \qquad (25)$$

$$\pi_{1jk} = \beta_{10k} + \beta_{11k}\text{SEX}_{jk} + \beta_{12k}\text{BLACK}_{jk} + \beta_{13k}\text{HISPANIC}_{jk} + r_{1jk} \qquad (26)$$

$$\beta_{00k} = \gamma_{000} + u_{00k} \qquad (27)$$

$$\beta_{pqk} = \gamma_{100} \qquad (28)$$

for $p = 0, 1$, $q = 1, 2, 3$, and $p = q \neq 0$.

Models with student level variables (GENDER, BLACK and HISPANIC) and one school level variables, percentage of minority students in a school, PMINORITY:

PP-4 Model:

$$\begin{aligned}\text{READ96}_{ij} &= \beta_{0j} + \beta_{1j}\text{READ95}_{ij} + \beta_{2j}\text{SEX}_{ij} + \beta_{3j}\text{BLACK}_{ij} + \beta_{4j}\text{HISPANIC}_{ij} \\ &\quad + \beta_{5j}\text{SEX}_{ij} \cdot \text{READ95}_{ij} + \beta_{6j}\text{BLACK}_{ij} \cdot \text{READ95}_{ij} \\ &\quad + \beta_{7j}\text{HISPANIC}_{ij} \cdot \text{READ95}_{ij} + r_{ij} \qquad (29)\end{aligned}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{PMINORITY}_j + u_{0j} \qquad (30)$$

$$\beta_{qj} = \gamma_{q0} \qquad (31)$$

for $q = 1, 2, \ldots, 7$.

Gain-4 Model:

$$\text{READGAIN9596}_{ijk} = \pi_{0jk} + \pi_{1jk}\text{READ95}_{ijk} + e_{ijk} \tag{32}$$

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}\text{SEX}_{jk} + \beta_{02k}\text{BLACK}_{jk} + \beta_{03k}\text{HISPANIC}_{jk} + r_{0jk} \tag{33}$$
$$\pi_{1jk} = \beta_{10k} + \beta_{11k}\text{SEX}_{jk} + \beta_{12k}\text{BLACK}_{jk} + \beta_{13k}\text{HISPANIC}_{jk} + r_{1jk} \tag{34}$$

$$\beta_{00k} = \gamma_{000} + \gamma_{001}\text{PMINORITY}_k + u_{00k} \tag{35}$$
$$\beta_{pqk} = \gamma_{100} \tag{36}$$

for $p = 0, 1$, $q = 1, 2, 3$, and $p = q \neq 0$.

For models PP-1, PP-2, PP-3 and PP-4:

$$i = 1, 2, \ldots, 5197 \text{ and } j = 1, 2, \ldots, 118.$$

Ranking of schools are obtained from $\hat{u_{0j}}$ which measures the deviation of the individual school's intercept from the overall intercept of the schools. This estimate is shrinkage adjusted by the number of repeated observations used to calculate this measure.

For models GAIN-1, GAIN-2, GAIN-3 and GAIN-4:

$$i = 1, 2, \ldots, 5197, \quad j = 1, 2, \ldots, 5197 \text{ and } k = 1, 2, \ldots, 118.$$

Ranking of schools are obtained from $\hat{u_{00k}}$ which measures the deviation of the individual school's intercept from the overall intercept of the schools. This estimate is also shrinkage adjusted by the number of repeated observations used to calculate this measure.

Classroom Effective Indices are obtained from the student residuals, namely r_{ij}, for the PP models and r_{ijk}, for the Gain Models, which measure the deviation of the students score from it's predicted value.

Results

School Effective Indices

Under the model assumptions, the u_{oj}'s for the 2-level models and u_{ook}'s for the 3-level models are normally distributed with a mean of zero. This fact is necessary if the school rankings obtained from the models are used to rank schools, and specially if the top schools are given performance awards. The need may arise to calculate confidence intervals and the normality assumtions will be vital. The descriptive statistics for the measures of School Effective Indices are as follow:

Model	Measure	Mean	St. Dev.	Kurtosis	Skewness	K-S P-Value
PP-1	u_{0j}	0.0000017	0.866	0.148	0.170	0.9126
Gain-1	u_{00k}	0.0052119	4.008	0.156	0.207	0.9546
PP-2	u_{0j}	-0.000002	0.703	0.907	0.442	0.9729
Gain-2	u_{00k}	0.0001864	3.135	0.649	0.454	0.5278
PP-3	u_{0j}	-0.000003	0.788	0.404	0.273	0.8197
Gain-3	u_{00k}	0.0018516	3.679	0.311	0.288	0.7969
PP-4	u_{0j}	-0.000000	0.696	0.102	0.534	0.9676
Gain-4	u_{00k}	0.0009754	3.180	0.828	0.522	0.8752

Kolmogorov-Smirnov Goodness of fit-test was carried out to determine if these measure are Normally distributed. The table above indicates K-S P-Values which are extremely high, hence failing to reject the Null hypothesis that the values are from a normal distribution. The values of Kurtosis for the models PP-2, Gain-2 and Gain-4 may suggest that the distribution is not normal, but the Kolmogorov-Smirnov Test proves otherwise.

The table below gives the correlations of the ranking measures for each model with one another and two school level variables MOBILITY(MOB) and PMINORITY(PMIN). PMINORITY is included in the models 2 and 4 to demonstrate how efficiently these models remove school level biases, and MOBILITY is included as a control.

	PP-1	G-1	PP-2	G-2	PP-3	G-3	PP-4	G-4	PMIN	MOB
PP-1	1.000									
G-1	0.988	1.000								
PP-2	0.885	0.869	1.000							
G-2	0.874	0.886	0.982	1.000						
PP-3	0.987	0.974	0.925	0.912	1.000					
G-3	0.978	0.989	0.906	0.922	0.987	1.000				
PP-4	0.882	0.865	0.996	0.979	0.932	0.912	1.000			
G-4	0.881	0.891	0.982	0.996	0.927	0.935	0.985	1.000		
PMIN	-0.459	-0.465	0.000	-0.012	-0.357	-0.374	-0.000	-0.028	1.000	
MOB	-0.215	0.202	-0.137	-0.112	-0.201	-0.188	-0.139	-0.128	0.246	1.000

From the table above we can observe that the School Rankings obtained from all the models are highly correlated with each other. The correlation range in value from 0.9819 to 0.9883. The correlations in the full model, with both student level and school level variables is 0.9847. Thus we can conclude that each pair of models, for the four different models considered, produces school rankings that are not significantly different from each other.

The last two columns on the table have two school level variables. Of these, PMINORITY was included in the Models PP-2, G-2 and PP-4, G-4. Comparing the correlations, we can conclude that any bias introduced by PMINORITY into the school ranking is removed by including this variable into the model. The Pre-test/Post-test models did a better task of removing the bias than the Gain model since the later models have higher correlations than

BEST COPY AVAILABLE

the former models. The school level variable MOBILITY was not included in any of the models and, as can be seen in the table above, it's bias remains in the school rankings.

Classroom Effective Indices

The process of obtaining Classroom Effective Indices from these models is more complicated than School Effective Indices. All the models above produce regression lines which are school specific, i.e. there is a regression line for each school. If we estimate $\hat{r_{ij}}$ for the PP models and $\hat{r_{ijk}}$ for the GAIN models using these lines we obtain student residuals within a specific school. Hence we calculate an average regression line(District Line) and calculate the student residuals from this line with respect to each model as follows.
For the Pre-Test/Post-Test models:

$$\hat{r_{ij}} = \text{READ96}_{ij} - (\gamma_{00} + \gamma_{10}\text{READ95}_{ij} + \gamma_{20}\text{SEX}_{ij} + \cdots) \tag{37}$$

For the GAIN models:

$$\hat{r_{ijk}} = \text{READ96}_{ijk} - (\gamma_{000} + \gamma_{100}\text{READ95}_{ijk}) \tag{38}$$

The following table illustrates the correlations of the student residuals obtained from the eight models.

	PP-1	G-1	PP-2	G-2	PP-3	G-3	PP-4	G-4
PP-1	1.000							
G-1	0.957	1.000						
PP-2	1.000	0.957	1.000					
G-2	0.957	1.000	0.957	1.000				
PP-3	0.997	0.955	0.996	0.955	1.000			
G-3	0.957	1.000	0.957	1.000	0.955	1.000		
PP-4	0.998	0.056	0.998	0.956	1.000	0.956	1.000	
G-4	0.957	1.000	0.957	1.000	0.955	1.000	0.956	1.000

As can be seen from the above table the Pre-Test/Post-Test Models(PP) and the Gain Models(G) are also highly correlated, with the lowest correlation having a $\rho = 0.955$. We can conclude from the above that the two types of models calculate district-wide residuals with significantly the same distribution.

There is major draw back of calculating district wide residuals from a multi-level model with schools as a level. Within a school, the student residuals, i.e., r_{ij} for j fixed has a normal distribution with a mean of zero. This is not the same for calculating student residuals district wide. Even if the model is not a multi-level model, but a linear regression model, the student residuals will not be normal. This fact should be considered when these residuals are analyzed and utilized.

Though the residuals of the various pairs of models are significantly correlated, this may not be evident from the descriptive statistics below.

Model	Mean	Median	St. Dev.	Kurtosis	Skewness	K-S P-Value
PP-1	-0.029	0.102	3.848	-0.171	-0.130	0.0026
Gain-1	-0.181	-1.857	19.43	0.537	0.372	0.0000
PP-2	-0.030	0.105	3.850	-0.181	-0.129	0.0015
Gain-2	-0.047	-1.619	19.4434	0.536	0.377	0.0000
PP-3	-0.017	0.081	3.826	-0.15	-0.120	0.0310
Gain-3	-0.087	-1.615	19.43	0.535	0.379	0.0000
PP-4	-0.021	0.116	3.830	-0.165	-0.123	0.0132
Gain-4	-0.130	-1.543	19.44	0.534	0.383	0.0000

Kolmogorov-Smirnov Goodness of fit-test was carried out to determine if these measure are Normally distributed. The table above indicates K-S P-Values which are extremely low, hence rejecting the null hypothesis that the values are from a normal distribution. The values of Kurtosis and Skewness for the models PP is significantly lower than that for the GAIN models. The medians for the PP models are closer to the mean than that of the GAIN models. The fact that the means of the residuals are close to zero is an encouragement. These analysis were carried out with 5197 student residuals.

By aggregating these student residuals by each classroom of the student, Classroom Effective Indices can be obtained.

Summary

Should we model Post-Test scores or Gains with Pre-Test scores as predictors in calculating student growth and school efficacy? This paper arrives at the conclusion that both types of models produce very similar results. Since the two types of models are comparable, the question arises which one should be used. If HLM software from SSI Inc.'s is used, then Pre-Test/Post-Test : two level models are more convenient and efficient than the Gain : three level models. In the two level models, more level-1 and level-2 variables can be introduced with proper centering to obtain complex models without any biases to the conditioning variables. This is not the same for 3-level HLM models. The models are very sensitive to multicollinearity and low variances in conditioning variables. As a result the number of level-2(student level characteristics) and level-3(school level) variables that can be introduced into the model is limited. In the models analyzed in this paper, the school level variable MOBILITY could not be introduced in the 3-level models due to software limitations, though MOBILITY could easily be entered into the 2-level models. Further, the iterative process of obtaining solutions to these 3-level hierarchical models using the EM algorithm has difficulties in calculating starting values to begin the iterative process. These starting values have to be cleverly estimated and placed into the model to start the iterative process.

All these difficulties with the 3-level models without any substantially different results that are obtained leads to the conclusion that the Pre-Test/Post-Test 2-level model is preferred over the 3-level Gain model.

References:

Anderson, T. W., (1971). *An Introduction to Multivariate Statistical Analysis.* New York: Wiley.

Bickel, P. J. and Doksum, K. A., (1977). *Mathematical Statistics.* Prentice Hall, Englewood Cliffs, New Jersey.

Bock, R. Darrel., (1989). *Multilevel Analysis of Educational Data.* San Diego, California: Academic Press.

Bryk, S. A. and Raudenbush, S. W., (1992). *Hierarchical Linear Models. Application and Data Analysis Methods.* Newburg Pass, California, Sage Publications.

Bryk, S. A. and Thum, Y. M., (1996). *Assessing the Productivity of Chicago Schools under Reform.* 1996 NORC Professional Development Training Session on Data Analysis: Concepts and Applications, Chicago.

Goldstein, H., (1987). *Multilevel Models in Educational and Social Research.* New York, New York: Oxford University Press.